Editorial Commentary

Whole genome sequencing in *Mycobacterium tuberculosis*

Andrea Spitaleri, Arash Ghodousi, Paolo Miotto, Daniela Maria Cirillo

Emerging Bacterial Pathogens Unit, Division of Immunology, Transplantation and Infectious Diseases, IRCCS San Raffaele Scientific Institute, Milan, Italy

Correspondence to: Daniela Maria Cirillo. Emerging Bacterial Pathogens Unit, Division of Immunology, Transplantation and Infectious Diseases, IRCCS San Raffaele Scientific Institute, via Olgettina 58, 20132 Milan, Italy. Email: cirillo.daniela@hsr.it.

Provenance: This is an invited article commissioned by the Guest Section Editor Dr. Maoshui Wang (Department of Lab Medicine, Shandong Provincial Chest Hospital, Jinan, China).

Comment on: Gygli SM, Keller PM, Ballif M, *et al.* Whole-Genome Sequencing for Drug Resistance Profile Prediction in Mycobacterium tuberculosis. Antimicrob Agents Chemother 2019;63. doi: 10.1128/AAC.02175-18.

Submitted Jun 19, 2019. Accepted for publication Jul 03, 2019.
doi: 10.21037/atm.2019.07.28
View this article at: http://dx.doi.org/10.21037/atm.2019.07.28

Fighting the emergence of multidrug-resistant (MDR) and extensively drug-resistant (XDR) tuberculosis (TB) requires a proper and personalized use of new antitubercular drugs. To achieve this goal, the sensitivity profile to antitubercular drug should be available in a short-time and results should be highly reliable be used for patient-care. However, among the main challenges affecting the performance of phenotypic drug susceptibility tests (DST) for *Mycobacterium tuberculosis complex* (*MTBC*), in addition to the long turnaround time, is the uncertainty on some of the critical concentrations (CCs) used for phenotypic DST. For lack of available and reliable data, those were established based on expert consensus and not on epidemiological cut-off values [abbreviated epidemiological cut-off (ECOFFs) or epidemiological cut-off values (ECV)], defined as it is the highest minimum inhibitory concentration (MIC) of organisms lacking phenotypically expressed resistance, e.g., the concentration of a drug that separate bacterial populations into those representative of a wild type population, and those with acquired or mutational resistance to the drug. The scarcity of data makes on MIC distributions of wild-type and mutants strains difficult to interpret in some cases and some of the established CCs cut the distribution of wild type and mutants (1). Whole genome sequencing (WGS), a relatively novel approach for *MTBC* strains analysis, has the great potential to provide fast and highly standardized DST results for prompt clinical guidance. Available published data and the extensively work carried out by the CRyPTIC project indicate that WGS performance for the detection of resistance is high for the two most important first-line drugs, e.g., rifampicin and isoniazid respectively, whereas the correlation between genomic WGS and the phenotypic DST data becomes smaller for other first-line drugs, e.g., ethambutol and pyrazinamide (2). In this context, the Authors of the manuscript by Gygli et al. compared WGS-based drug resistance profiles with culture-based quantitative DST methods (namely the 7H10 agar proportion and the MGIT 960) for eleven drugs among a random sample of patients, which should be representative of the geographically-defined population in the high and low burden TB countries (3). Although the sample size presented in this manuscript is very limited (189 isolated strains) and the geographical representativeness is low for certain lineages (ancient lineages L1, L5 and L6 account all together for 8 isolates in the study), the Authors were able to underline the opportunity offered by WGS analysis in providing reliable information for the majority of the drugs in use for treatment. At the same time, the study highlights the challenges in providing genotype-phenotype correlations for clinical management of TB.

The authors show that the MIC distributions of wild-type and mutant strains are well separated for most of the first and second-line drugs, allowing the possibility to clearly define the related ECOFFs (rifampicin, rifabutin, isoniazid, kanamycin, amikacin, capreomycin, streptomycin, pyrazinamide). For isoniazid, rifampicin/rifabutin, amikacin and streptomycin, the Authors suggest the introduction of "clinical breakpoints" helpful to define when MIC

increases linked to specific mutations is still within the therapeutic range of the drug. It needs to be considered that in some cases resistance emerge as low level and then progresses to more effective (high-level) resistance (4). Despite these limitations and uncertainties remain to be addressed as insufficient evidence of the safety and overall therapeutic outcomes exist, there is an increasing interest in the possibility to overcome low-level resistance by the use of higher drug dosage (5,6). Current TB treatment guidelines refer only to the possibility to use isoniazid or moxifloxacin at higher doses under specific conditions (7), however the data from Gygli et al. can contribute in better understanding the correlation between genotype, phenotype and pharmacokinetic/dynamic data. For ethambutol, moxifloxacin, ethionamide the Authors showed overlapping MIC distribution of the wild-type and mutant populations. On the one hand, the overlap in the MIC distribution in moxifloxacin-resistant isolates is only partial and with 80% of sensitivity, on the other, in case of ethambutol and ethionamide the overlapping is explained by a large number of polymorphisms in resistance-conferring genes. Noteworthy, in case of ethambutol the distribution curves of wild-type and mutants in embB resistance gene, could be better separated by adjusting the critical concentration at a lower value, from 5 to 2.5 mg/L. Importantly, the Authors reported that some mutations cause high variability in the increase of the MIC conferred by an identical mutation. This can have clinically relevant implications when there is a significant overlap between the MICs of the mutant and wild-type strains, making the classification of these isolates difficult to achieve by current WGS approaches.

Overall, the mutations identified by WGS have high accuracy in classifying as susceptible/resistant clinical isolates. However, for quantifying the level of resistance, which is not always possible by genotypic DST alone, MIC testing can be useful (1). MIC testing can resolve discrepancies between genotypic and phenotypic DST results for resistance mechanisms with MIC distributions that are divided by a clinical breakpoint that causes poor reproducibility for phenotypic DST even if the clinical breakpoint is greater or equal to the epidemiological cut-off value (8). There are other reasons which may explain the discrepancy between WGS data and phenotypic DST. A first reason could be related to the detection of low-level antimicrobial drug resistances that are difficult to characterize by DST methods due to the intrinsic low sensitivity, whereas WGS could in principle detect these low-level mutations and be more accurate in predicting drug susceptibility with respect DST approaches (9). Second, the current WGS-based drug resistance predictions rely only the analysis of the few genes known to be involved in drug resistance. In general, mutations resulting in antimicrobial resistance alter the antibiotic action via different mechanisms, such as modifications of the antimicrobial target, a decrease in the drug uptake, activation of efflux mechanisms to extrude the drug, and finally via global changes in important metabolic pathways via modulation of regulatory networks. An extensive knowledge of the mechanisms should include extended pathways rather than focusing on a small number of genes directly involved in drug interactions (e.g., drug activation, drug modification, or drug target).

MTBC can adopt several mechanisms of resistance to survive the effect of the antibiotic through multiple biochemical pathways in which different genes are involved. Thus, resistance mechanisms arising due to acquired mutational changes may be the result of several concomitant causes.

Highly conserved and essential genes, e.g., *rpoB*, *rpsL* and *rrs*, have a very good correlation with the phenotypic DST, showing that WGS is undoubtedly a powerful approach able to discriminate susceptible against resistant strains. Other not essential genes, e.g., *pncA* and *ethA*, involved in the conversion of prodrugs into their active forms, e.g., pyrazinamide and ethionamide, present several not lethal mutations for the *MTBC*, leading to a substantial overlap along the MIC distribution. Finally, genes that show also high degree of polymorphism, e.g., *embB*, are difficult to interpret as function of the MIC distribution. An explanation of this results, is that essential genes are also the end target for the antibiotic, whereas the not essential genes are often not directly involved in the drug interaction.

Noteworthy, mutations causing changes in expression profiling are poorly represented among the current genetic markers of resistance. At this end, the role of mutations causing changes in gene expression because the disruption of regulatory pathways remains completely unexplored. Interestingly, also synonymous mutations have found to affect this complex regulatory network (10). Similarly, the role of mutations in intergenic regions mapping in putative non-coding regulatory small RNAs remains unknown and underexplored (11). Third, standard pipelines for the differentiation of mutations, including phylogenetic and synonymous mutations, in correlation with drug resistance are still missing or lack a full agreement among all stakeholders. Fourth, the discordance between

phenotypic and WGS-based DST could be related to the presence of mixed *MTBC* infections (TB disease caused by more than one distinct *MTBC* strain) or heterogeneous resistance (mixture of wild type and mutated population at same location) (12). These aspects were not taken into consideration by the Authors. Although WGS based DST analysis using heterozygous base calling could provide better resolution of mixed/heterozygous infections, clear thresholds still need to be established (13,14).

In addition, most of the currently available pipelines cannot yet reliably identify heterogeneous Indels (mixture of wild type and Indels at same location) in presence of minor sub-populations (15). This may be a challenge to accurately predict phenotype from sequence information. Moreover, the clinical relevance of such minor populations will need to be evaluated for each drug taking into account the global resistance pattern, and the treatment regimen employed.

In summary, the results shown in this manuscript highlights the strength and weakness of the WGS applied in the clinical routine for *MTBC* diagnosis. Efforts from multiple research groups, together with global stakeholders, are needed to further improve and integrate *MTBC* genomics into healthcare and surveillance programs. Nowadays, large efforts in TB drug discovery are in place to reduce the duration of the treatment for both drug-susceptible and drug-resistant TB, and to improve cure rates. Many novel anti-TB drug candidate with novel mechanism of actions are in pre-clinical development. Bedaquiline and Delamanid have received provisional approval by Food and Drug Administration (FDA) and European Medicine Agency (EMA) since 2012 and represent the front-line treatment for MDR and XDR-TB regimen (8). Efforts to evaluate their relationship between DST and WGS analysis is ongoing in order to find resistance conferring genes (16). In order to be able to properly interpret rare mechanisms of resistance, it is essential that large international Consortia such as CRyPTIC and RESEQ-TB will succeed in collecting and analyze a huge number of genomic, phenotypic and clinical data from strains collected at the global level.

Acknowledgments

None.

Footnote

Conflicts of Interest: The authors have no conflicts of interest to declare.

Ethical Statement: The authors are accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

References

1. Ghodousi A, Tagliani E, Karunaratne E, et al. Isoniazid resistance in Mycobacterium tuberculosis is a heterogeneous phenotype comprised of overlapping MIC distributions with different underlying resistance mechanisms. Antimicrob Agents Chemother 2019. doi: 10.1128/AAC.00092-19
2. CRyPTIC Consortium and the 100,000 Genomes Project, Allix-Béguec C, Arandjelovic I, et al. Prediction of Susceptibility to First-Line Tuberculosis Drugs by DNA Sequencing. N Engl J Med 2018;379:1403-15.
3. Gygli SM, Keller PM, Ballif M, et al. Whole-Genome Sequencing for Drug Resistance Profile Prediction in Mycobacterium tuberculosis. Antimicrob Agents Chemother 2019;63. doi: 10.1128/AAC.02175-18.
4. Baquero F. Low-level antibacterial resistance: a gateway to clinical resistance. Drug Resist Updat 2001;4:93-105.
5. Xu Y, Wu J, Liao S, et al. Treating tuberculosis with high doses of anti-TB drugs: mechanisms and outcomes. Ann Clin Microbiol Antimicrob 2017;16:67.
6. Goutelle S, Bourguignon L, Maire P, et al. The Case for Using Higher Doses of First Line Anti-Tuberculosis Drugs to Optimize Efficacy. Curr Pharm Des 2014;20:6191-206.
7. WHO consolidated guidelines on drug-resistant tuberculosis treatment. Geneva: World Health Organization, 2019.
8. World Health Organization. Technical report on critical concentrations for drug susceptibility testing of medicines used in the treatment of drug- resistant tuberculosis. Geneva: World Health Organization, 2018.
9. Horne DJ, Pinto LM, Arentz M, et al. Diagnostic accuracy and reproducibility of WHO-endorsed phenotypic drug susceptibility testing methods for first-line and second-line antituberculosis drugs. J Clin Microbiol 2013;51:393-401.
10. Jajou R, van der Laan T, de Zwaan R, et al. WGS more accurately predicts susceptibility of Mycobacterium tuberculosis to first-line drugs than phenotypic testing. J Antimicrob Chemother 2019. doi: 10.1093/jac/dkz215.
11. Miotto P, Forti F, Ambrosi A, et al. Genome-wide discovery of small RNAs in Mycobacterium tuberculosis.

PLoS One 2012;7:e51950.
12. Noto JM, Piazuelo MB, Chaturvedi R, et al. Strain-specific suppression of microRNA-320 by carcinogenic Helicobacter pylori promotes expression of the antiapoptotic protein Mcl-1. Am J Physiol Gastrointest Liver Physiol 2013;305:G786-96.
13. Iketleng T, Lessells R, Dlamini MT, et al. Mycobacterium tuberculosis Next-Generation Whole Genome Sequencing: Opportunities and Challenges. Tuberc Res Treat 2018;2018:1298542.
14. Hatherell HA, Colijn C, Stagg HR, Jackson C, Winter JR, Abubakar I. Interpreting whole genome sequencing for investigating tuberculosis transmission: a systematic review. BMC Med 2016;14:21.
15. Nimmo C, Shaw LP, Doyle R, et al. Whole genome sequencing Mycobacterium tuberculosis directly from sputum identifies more genetic diversity than sequencing from culture. BMC Genomics 2019;20:389.
16. Ezewudo M, Borens A, Chiner-Oms Á, et al. Integrating standardized whole genome sequence analysis with a global Mycobacterium tuberculosis antibiotic resistance knowledgebase. Sci Rep 2018;8:15382.

Cite this article as: Spitaleri A, Ghodousi A, Miotto P, Cirillo DM. Whole genome sequencing in *Mycobacterium tuberculosis*. Ann Transl Med 2019;7(Suppl 6):S197. doi: 10.21037/atm.2019.07.28

